

RAW SEQUENCE LISTING

**The Biotechnology Systems Branch of the Scientific and Technical
Information Center (STIC) no errors detected.**

Application Serial Number: 09/664,958A
Source: IFW16
Date Processed by STIC: 10/21/05

ENTERED



IFW

IFW16

RAW SEQUENCE LISTING

DATE: 10/21/2005

PATENT APPLICATION: US/09/664,958A

TIME: 09:05:44

Input Set : A:\60240 SeqList 10.14.05.ST25.txt

Output Set: N:\CRF4\10212005\I664958A.raw

3 <110> APPLICANT: The Trustees of Columbia University in the City of New York
 4 Trakht, Ilya
 5 Canfield, Robert
 6 Kalantarov, Gary
 7 Rudchenko, Sergei
 9 <120> TITLE OF INVENTION: Novel Tumor-Associated Marker
 11 <130> FILE REFERENCE: 0575/60240/JPW/AJM/AJD
 13 <140> CURRENT APPLICATION NUMBER: 09/664,958A
 -> 14 <141> CURRENT FILING DATE: 2000-09-18
 16 <160> NUMBER OF SEQ ID NOS: 53
 18 <170> SOFTWARE: PatentIn version 3.3
 20 <210> SEQ ID NO: 1
 21 <211> LENGTH: 333
 22 <212> TYPE: PRT
 23 <213> ORGANISM: Human
 25 <400> SEQUENCE: 1
 27 Met Pro Leu Gly Leu Gly Arg Arg Lys Lys Ala Pro Pro Leu Val Glu
 28 1 5 10 15
 31 Asn Glu Glu Ala Glu Pro Gly Arg Gly Gly Leu Gly Val Gly Glu Pro
 32 20 25 30
 35 Gly Pro Leu Gly Gly Gly Gly Ser Gly Gly Pro Gln Met Gly Leu Pro
 36 35 40 45
 39 Pro Pro Pro Pro Ala Leu Arg Pro Arg Leu Val Phe His Thr Gln Leu
 40 50 55 60
 43 Ala His Gly Ser Pro Thr Gly Arg Ile Glu Gly Phe Thr Asn Val Lys
 44 65 70 75 80
 47 Glu Leu Tyr Gly Lys Ile Ala Glu Ala Phe Arg Leu Pro Thr Ala Glu
 48 85 90 95
 51 Val Met Phe Cys Thr Leu Asn Thr His Lys Val Asp Met Asp Lys Leu
 52 100 105 110
 55 Leu Gly Gly Gln Ile Gly Leu Glu Asp Phe Ile Phe Ala His Val Lys
 56 115 120 125
 59 Gly Gln Arg Lys Glu Val Glu Val Phe Lys Ser Glu Asp Ala Leu Gly
 60 130 135 140
 63 Leu Thr Ile Thr Asp Asn Gly Ala Gly Tyr Ala Phe Ile Lys Arg Ile
 64 145 150 155 160
 67 Lys Glu Gly Ser Val Ile Asp His Ile His Leu Ile Ser Val Gly Asp
 68 165 170 175
 71 Met Ile Glu Ala Ile Asn Gly Gln Ser Leu Leu Gly Cys Arg His Tyr
 72 180 185 190
 75 Glu Val Ala Arg Leu Leu Lys Glu Leu Pro Arg Gly Arg Thr Phe Thr
 76 195 200 205
 79 Leu Lys Leu Thr Glu Pro Arg Lys Ala Phe Asp Met Ile Ser Gln Arg

p.6

RAW SEQUENCE LISTING

DATE: 10/21/2005

PATENT APPLICATION: US/09/664,958A

TIME: 09:05:44

Input Set : A:\60240 SeqList 10.14.05.ST25.txt

Output Set: N:\CRF4\10212005\I664958A.raw

```

80      210      215      220
83 Ser Ala Gly Gly Arg Pro Gly Ser Gly Pro Gln Leu Gly Thr Gly Arg
84 225      230      235      240
87 Gly Thr Leu Arg Leu Arg Ser Arg Gly Pro Ala Thr Val Glu Asp Leu
88      245      250      255
91 Pro Ser Ala Phe Glu Glu Lys Ala Ile Glu Lys Val Asp Asp Leu Leu
92      260      265      270
95 Glu Ser Tyr Met Gly Ile Arg Asp Thr Glu Leu Ala Ala Thr Met Val
96      275      280      285
99 Glu Leu Gly Lys Asp Lys Arg Asn Pro Asp Glu Leu Ala Glu Ala Leu
100      290      295      300
103 Asp Glu Arg Leu Gly Asp Phe Ala Phe Pro Asp Glu Phe Val Phe Asp
104 305      310      315      320
107 Val Trp Gly Ala Ile Gly Asp Ala Lys Val Gly Arg Tyr
108      325      330
111 <210> SEQ ID NO: 2
112 <211> LENGTH: 720
113 <212> TYPE: DNA
114 <213> ORGANISM: Human
116 <400> SEQUENCE: 2
117 cacggggagg cggaggcagc ggcggcgccg gcggcgccgg cggcgccggc ggagcagatc 60
119 ttctggtgac cccacttctc gctgctcatg ccgctgggac tggggcgccg gaaaaaggcg 120
121 cccctctag tggaaaatga ggaggctgag ccaggccgtg gagggctggg cgtgggggag 180
123 ccagggcctt tgggcggagg tgggtcgggg ggcccccaa tgggcttgcc cccctcccc 240
125 ccagccctgc ggccccgcct tgtgttcac accagctgg cccatggcag tccactggc 300
127 cgcacgagg gggtcaccaa cgtcaaggag ctgtatggca agattgccga ggccttccgc 360
129 ctgccaaactg ccgaggtgat gttttgcacc ctgaacaccc acaaagtggg catggacaag 420
131 ctcttggggg gccaaatcgg gctggaggac ttcatcttcg ccacgtgaa ggggcagcgc 480
133 aaggaggtgg aggtgttcaa gtcggaggat gcactcgggc tcaccatcac ggacaacggg 540
135 gctggctacg ccttcatcaa gcgcatcaag gagggcagcg tgatcgacca catccacctc 600
137 atcagcgtgg gcgacatgat cgaggccatt aacgggcaga gcctgctggg ctgccggcac 660
139 tacgaagtgg cccggtgct caaggaactg ccccgaggcc gtaccttcac gctgaagctc 720
142 <210> SEQ ID NO: 3
143 <211> LENGTH: 9
144 <212> TYPE: PRT
145 <213> ORGANISM: Human
147 <400> SEQUENCE: 3
149 Lys Leu Leu Gly Gly Gln Ile Gly Leu
150 1 5
153 <210> SEQ ID NO: 4
154 <211> LENGTH: 10
155 <212> TYPE: PRT
156 <213> ORGANISM: Human
158 <400> SEQUENCE: 4
160 Ser Leu Leu Gly Cys Arg His Tyr Glu Val
161 1 5 10
164 <210> SEQ ID NO: 5
165 <211> LENGTH: 6263
166 <212> TYPE: DNA

```

RAW SEQUENCE LISTING

DATE: 10/21/2005

PATENT APPLICATION: US/09/664,958A

TIME: 09:05:44

Input Set : A:\60240 SeqList 10.14.05.ST25.txt

Output Set: N:\CRF4\10212005\I664958A.raw

```

167 <213> ORGANISM: Human
169 <400> SEQUENCE: 5
170 catcagcggg cgggggtgtc gccgaacagg ctgtctcgca gagcccgccg cgaccccgcg      60
172 ccgccccgcc ccgcggcctg cctgccagag gagccgaggg ggccgcccct cgcccaacct      120
174 gcccagacatg gggaaacccg ggcccaggcg tgtgtgtcac catgacaaca gagacaggcc      180
176 ccgactctga ggtgaagaaa gctcaggagg agggcccgca gcagcccgag gctgtgtccg      240
178 ctgtgaccac ccctgtgacc cctgcaggcc acggccaccc agaggccaac tccaatgaga      300
180 agcatccatc ccagcaggac acgcggcctg ctgaacagag cctagacatg gaggagaagg      360
182 actacagtga ggccgatggc ctttcggaga ggaccacgcc cagcaaggcc cagaaatcgc      420
184 cccagaagat tgccaagaaa tacaagagtg ccatctgccg ggtcactctg cttgatgcct      480
186 cggagtatga gtgtgagggt gagaaacatg gccggggcca ggtgtgtttt gacctggtct      540
188 gtgaacacct caacctccta gagaaggact acttcggcct gaccttctgt gatgtgaca      600
190 gccagaagaa ctggctggac ccctccaagg agatcaagaa gcagatccgg agtagccct      660
192 ggaattttgc cttcacagtc aagttctacc cgctgatcc tgcccagctg acagaagaca      720
194 tcacaagata ctacctgtgc ctgcagctgc gggcagacat catcacgggc cggctgccat      780
196 gtccttttgt cacgcatgcc ctactgggct cctacgctgt gcaggctgag ctgggtgact      840
198 atgatgctga ggagcatgtg ggcaactatg tcagcgagct ccgcttcgcc cctaaccaga      900
200 cccgggagct ggaggagagg atcatggagc tgcataagac atataggggg atgaccccg      960
202 gagaagcaga aatccacttc ttagagaatg ccaagaagct ttccatgtac ggagtagacc      1020
204 tgcaccatgc caaggactct gagggcacatg acatcatgtt aggcgtttgt gccaatggcc      1080
206 tgtcatctta ccgggaccgg ctgagaatca accgctttgc ctggcccaag atcctcaaga      1140
208 tctcctacaa gaggagtaac ttctatatca agatccggcc tggggagtat gagcaatttg      1200
210 agagcacaat tggttttaag ctcccaaacc accggtcagc caagagactg tggaaggctct      1260
212 gcatcgagca tcatacattc ttccggctgg tgtccctga gccccaccc aagggttcc      1320
214 tgggtgatgg cccaagttc cggtagctg ggagagccca ggcacagact cgccaggcca      1380
216 gcgccctcat tgaccggcct gcaccttct ttgagcgttc ttccagcaaa cggtagacca      1440
218 tgtcccgtag ccttggtgga gcagagttct cccgcccagc ctcggtcagc gagaacctatg      1500
220 atgcagggcc tgacggtgac aagcgggatg aggatggcga gtctgggggg caacggtcag      1560
222 aggtgagga gggagagggtc aggaactcaa ccaagatcaa ggagctaaag ccggagcagg      1620
224 aaaccacgcc gagacacaag caggagttct tagacaagcc agaagatgtc ttgctgaagc      1680
226 accaggccag catcaatgag ctcaaaagga ccctgaagga gcccaacagc aaactcatcc      1740
228 accgggatcg agactgggaa cgggagcgca ggctgcctc ctcccccgcc tccccctccc      1800
230 ccaagggcac ccctgagaaa gccaatgaga gagcagggtc gaggggagggc tccgaggaga      1860
232 aagtcaaacc accacgtccc cgggccccag agagtgcac aggcgatgag gaccaggacc      1920
234 aggagaggga cacggtgttc ctgaaggaca accacctggc cattgagcgc aagtgtcca      1980
236 gcatcacggt cagctctacg tctagcctgg aggtgaggt ggacttcacg gtcattggtg      2040
238 actacatgg cagcgccctc gaagacttct cccgcagcct gcctgagctc gaccgggaca      2100
240 aaagcgactc ggacactgag ggccctgtgt tctcccgga tctcaacaag ggggccccca      2160
242 gccaggatga tgagtctggg ggcattgagg acagcccgga tcgagggggc tgtccaccc      2220
244 cggatatgcc ccagtttgag cccgtgaaaa cagaaaccat gactgtcagc agtctggcca      2280
246 ttagaaagaa gattgagccg gaggccgtac tgcagaccag agtctccgct atggataaca      2340
248 cccagcaggt tgatgggagt gcctcagtgg ggagggagtt catagcaacc actccctcca      2400
250 tcaccacgga gaccatctc gaccacctgg agaacagttc caagtccggg aagggggcag      2460
252 ctgccatgat cccaggccca cagacggtgg ccacggaaat ccgttctctt tctccgatca      2520
254 tcgggaaaga tgtcctcacc agcacctacg gcgccactgc ggaaaccctc tcaacctcca      2580
256 ccaccaccca tgtcaccaaa actgtgaaag gagggttttc tgagacaagg atcgagaagc      2640
258 gaatcatcat tactgggat gaagatgtcg atcaagacca ggccctggct ttggccatca      2700
260 aggaggccaa actgcagcat cctgatatgc tggtaaccaa agctgtcgtc tacagagaaa      2760
262 cagacccatc cccagaggag agggacaaga agccacagga atcctgacct ctgtgaagag      2820

```

RAW SEQUENCE LISTING

DATE: 10/21/2005

PATENT APPLICATION: US/09/664,958A

TIME: 09:05:44

Input Set : A:\60240 SeqList 10.14.05.ST25.txt

Output Set: N:\CRF4\10212005\I664958A.raw

264	atcctggcat	ttctgggtcca	acccaagcca	gagaaccatt	aagaaggggc	cttcattctg	2880
266	gattctccga	cgcaacactg	acgtcccagc	tgcgacgtac	tgctactgat	gagagactgg	2940
268	gaagggaaaa	gcatatatat	atagatatat	agagatatag	atatatatac	aggaaacacc	3000
270	gcatccttgc	actgctgctg	gggctggcag	agcagttggc	tgacagcaac	aaccgacatc	3060
272	tgaacaccta	catttccttt	gcagacaaat	tgaagaactg	gtgggatttt	tttcaagaaa	3120
274	aaaaattata	taataactat	aatcccttgc	tcaccccttt	ccccgcgcaa	ataagaaacg	3180
276	caagccagac	cacgatgatt	gtagaagtcc	ctcccgcctt	ggttctgcac	gttacagtta	3240
278	gcagacgagc	aattccattt	gttcttctcc	agcatctcta	aggccactt	gaatgcaaag	3300
280	gaaaacactt	gcacagcaaa	gcaagagaag	tcacagcagc	aagacacgca	cagtcaacca	3360
282	ttttccgaga	aaaaaagaaa	attccccact	tggaaagaaa	gaggaggaac	actggattct	3420
284	tactttctgg	atcttgacac	tgggctgcaa	aacctacett	cctctctccc	gcctcccttc	3480
286	accctcaact	ctcaatgtct	tgtgtcatt	ttctgtctcg	gtccctctct	cccccttccc	3540
288	ccttccccca	ccccacaccc	ttcacctctt	gtgtcctggg	ccttctgagg	gccactgcag	3600
290	atgactctcc	tttgaaatga	gaaaaagaaa	agaaagcaag	aacagaaaac	gaagccacag	3660
292	gaaggggaagt	agacattgta	tgtttatggg	ttctcattat	gaaggtgcag	cttgtaggag	3720
294	gtttgtacgg	atgtgctttg	aagtattgta	tattacatat	aacaggaaaa	aatattaata	3780
296	aacagtgtctg	gtaagtatga	agctgacatt	ctaaaattat	aattatctga	ctgtgattga	3840
298	tgtatcctga	ggttccctaga	tctcactgaa	ctggcccagc	taaggagacc	tggactctgg	3900
300	gtgtgggttg	gtcacagta	ggggtgcag	gggtcagtg	agtaatactg	tgtgtgggtg	3960
302	ttgtaattgg	ttgattgggtg	gggaggggtg	gggggcccct	atggagaggt	gtgggttttg	4020
304	caagaaagaa	gcaacacaga	tgtcgtcccc	aaaatgccag	ttcaagacac	cttctccctg	4080
306	ccccctggg	agtaacagtc	agggcctggg	ctgtgctcag	gtactgggtc	ccagtctggg	4140
308	actctgctgc	tgaagtgtcc	acagttaggg	tccctggctt	agtccttata	tccctacggg	4200
310	gcttgccctg	gttttcagtc	ttctctctct	ttcttctctt	tttttttttt	tgccacattc	4260
312	tgcccttccc	tgacccctt	gtaataacca	actccatata	caaagggagg	tgggtgctct	4320
314	agccattgta	gaagatgggtg	gctttaacct	gactgtctaa	aaattcccag	ctaagccttt	4380
316	tctctacttc	tcttccctgt	tctgaatcat	ttcttcttct	caggccaaag	tagccatggg	4440
318	aaggaggctt	catggggcag	accctgaaag	atcaaaaactg	catttgcaaa	gccctccctt	4500
320	gtcccaggac	aaagctgaga	ctgacgggtg	atgttgctca	taggctccag	ctctgcataa	4560
322	gaccttggct	tggagacctc	cctctcagtc	aacagctgaa	ctctgagctt	gtgccagaa	4620
324	attaccccaa	gaccacagga	acccttcaag	aagctcccat	cacaagcttg	gcattgctct	4680
326	ctgccacacg	tgggcttcc	caggcttgct	tgccacaagc	tacttctctg	agctcagaaa	4740
328	gtgccccttg	atgagggaaa	atgtcccact	gcactgcgaa	tttctcagtt	ccattttacc	4800
330	tcccagtcct	ccttctaaac	cagttaataa	attcattcca	caagtattta	ctgattacct	4860
332	gcttgtgcca	gggactatct	tcaggctgaa	gaaggtggga	ggggagggcg	gaacctgagg	4920
334	agccacctga	gccagcttta	tatttcaacc	atggctggcc	catctgagag	catctcccca	4980
336	ctctcgccaa	cctatcgggg	catagcccag	ggatgcccc	aggcgccca	ggtagatgc	5040
338	gtccctttgg	cttgtcagtg	atgacataca	ccttagctgc	ttagctgggtg	ctggcctgag	5100
340	gcagggcagg	aatcagaat	agcatttgct	tctctgggca	aatgggaagt	tcagcggggc	5160
342	agcagaatca	gtggcattcc	ccctgggtgca	ggcgggtggg	tccactccaa	ctccccctga	5220
344	gtgtagcagc	acactttcca	tacaccaggt	tctttctaca	atcctgggtg	aaaagccaca	5280
346	gaaccttctt	cctgcccttc	ttgagagttc	ccccctcttc	tgggtcaaga	gctggagtgg	5340
348	tggctccatc	ctctctgggc	cacttcgggtc	taggaactca	tctttgcagg	aaccaggagt	5400
350	cctgagcaca	ctgaacacac	ctcagaggga	ggatccttgt	tgtggatttt	gcacctggct	5460
352	ttggggcagg	gggtgaagtga	ccaggcttag	cttgtggagt	ttatgggcca	ccagggtttg	5520
354	gggaaatcac	catcccgcg	atgctgtgac	ctcccttcta	cggagatgca	ggcagtgcc	5580
356	cgagggagga	ggggacctgc	aaagctagaa	tctagggcac	tgtttctctc	ccatccttct	5640
358	ctttgtagag	aatagagacg	tttgtcttgt	ctgtcttcaa	cctacttttc	cttttctctt	5700
360	ttttgtttct	catcctctct	gtgccacctc	tccacccagg	aggccatgta	gcatagtgga	5760

RAW SEQUENCE LISTING

DATE: 10/21/2005

PATENT APPLICATION: US/09/664,958A

TIME: 09:05:44

Input Set : A:\60240 SeqList 10.14.05.ST25.txt

Output Set: N:\CRF4\10212005\I664958A.raw

```

362 aaaagtcctt gagggcggtt aggagttctg ggtgaccatc ctggctcagc tcctaactca 5820
364 ccatgtgaca tcaggetatc cccattcccc ctcttggggc tcagtttccc gacttgcaaa 5880
366 ataagcagaa agaaccagat gctctccagg gtctttttct actttgctat ctcatgggtc 5940
368 ttcatTTTTt cttattttgt tttctctgga tcttttccat ctgagggtac aggaagtacc 6000
370 aggacctgtt tcagtttttg aatcctgcaa gcacattcca agactggcct gaaactgcat 6060
372 gagcaacatc actcgaaata attttttttt tcaaaagcac cttacaacc aattgcatg 6120
374 ctgtcctgtt cctttttact cacacccttc tctcctttct cgtccccatg ctccccacc 6180
376 tcagtgtctc gtgctgtatg cgtgtgctct ctgttcttgt atactcaata taagtgaat 6240
378 aaatgtgttt gatgctgaac cat 6263
381 <210> SEQ ID NO: 6
382 <211> LENGTH: 933
383 <212> TYPE: PRT
384 <213> ORGANISM: Human
386 <400> SEQUENCE: 6
388 Ser Ala Gly Gly Gly Val Ala Glu Gln Ala Ala Pro Gln Ser Pro Pro
389 1 5 10 15
392 Arg Pro Arg Ala Ala Pro Pro Arg Gly Leu Pro Ala Arg Gly Ala Glu
393 20 25 30
396 Gly Ala Ala Pro Arg Pro Thr Cys Pro Thr Trp Gly Thr Pro Gly Pro
397 35 40 45
400 Gly Val Leu Val Thr Met Thr Thr Glu Thr Gly Pro Asp Ser Glu Val
401 50 55 60
404 Lys Lys Ala Gln Glu Glu Ala Pro Gln Gln Pro Glu Ala Ala Ala Ala
405 65 70 75 80
408 Val Thr Thr Pro Val Thr Pro Ala Gly His Gly His Pro Glu Ala Asn
409 85 90 95
412 Ser Asn Glu Lys His Pro Ser Gln Asp Thr Arg Pro Ala Glu Gln Ser
413 100 105 110
416 Leu Asp Met Glu Glu Lys Asp Tyr Ser Glu Ala Asp Gly Leu Ser Glu
417 115 120 125
420 Arg Thr Thr Pro Ser Lys Ala Gln Lys Ser Pro Gln Lys Ile Ala Lys
421 130 135 140
424 Lys Tyr Lys Ser Ala Ile Cys Arg Val Thr Leu Leu Asp Ala Ser Glu
425 145 150 155 160
428 Tyr Glu Cys Glu Val Glu Lys His Gly Arg Gly Gln Val Leu Phe Asp
429 165 170 175
432 Leu Val Cys Glu His Leu Asn Leu Leu Glu Lys Asp Tyr Phe Gly Leu
433 180 185 190
436 Thr Phe Cys Asp Ala Asp Ser Gln Lys Asn Trp Leu Asp Pro Ser Lys
437 195 200 205
440 Glu Ile Lys Lys Gln Ile Arg Ser Ser Pro Trp Asn Phe Ala Phe Thr
441 210 215 220
444 Val Lys Phe Tyr Pro Pro Asp Pro Ala Gln Leu Thr Glu Asp Ile Thr
445 225 230 235 240
448 Arg Tyr Tyr Leu Cys Leu Gln Leu Arg Ala Asp Ile Ile Thr Gly Arg
449 245 250 255
452 Leu Pro Cys Ser Phe Val Thr His Ala Leu Leu Gly Ser Tyr Ala Val
453 260 265 270
456 Gln Ala Glu Leu Gly Asp Tyr Asp Ala Glu Glu His Val Gly Asn Tyr

```

RAW SEQUENCE LISTING ERROR SUMMARY DATE: 10/21/2005
PATENT APPLICATION: US/09/664,958A TIME: 09:05:45

Input Set : A:\60240 SeqList 10.14.05.ST25.txt
Output Set: N:\CRF4\10212005\I664958A.raw

Base Note:

a of n and/or Xaa have been detected in the Sequence Listing. Please review the
quence Listing to ensure that a corresponding explanation is presented in the <220>
<223> fields of each sequence which presents at least one n or Xaa.

1#:15; N Pos. 23,42,48,1105

1#:17; N Pos. 23,42

VERIFICATION SUMMARY

DATE: 10/21/2005

PATENT APPLICATION: US/09/664,958A

TIME: 09:05:45

Input Set : A:\60240 SeqList 10.14.05.ST25.txt

Output Set: N:\CRF4\10212005\I664958A.raw

14 M:271 C: Current Filing Date differs, Replaced Current Filing Date

1689 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:15 after pos.:0

341 Repeated in SeqNo=15

1810 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:17 after pos.:0